Culture-Fair Cognitive Ability Assessment

Information Processing and Psychophysiological Approaches

Steven P. Verney
University of New Mexico

Eric Granholm
University of California, San Diego, and
San Diego Veterans Affairs Healthcare System

Sandra P. Marshall
Vanessa L. Malcarne
Dennis P. Saccuzzo
San Diego State University

Valid assessment with diverse populations requires tools that are not influenced by cultural elements. This study investigated the relationships between culture, information processing efficiency, and general cognitive capacities in samples of Caucasian and Mexican American college students. Consistent with the neural efficiency hypothesis, pupillary responses (indexing mental effort) and detection accuracy scores on a visual backward-masking task were both significantly related to the Wechsler Adult Intelligence Scale–Revised (WAIS-R) Full Scale scores. These measures of information processing efficiency were similar in the two groups. However, they were related only to Caucasian American, but not to a comparable sample of Mexican American, students' WAIS-R scores. Therefore, the differential validity in prediction suggests that the WAIS-R test may contain cultural influences that reduce the validity of the WAIS-R as a measure of cognitive ability for Mexican American students. Information processing and psychophysiological approaches may be helpful in developing culture-fair cognitive ability measures.

Keywords: cognitive ability; information processing; intelligence; psychophysiology; pupillary response; racial and ethnic differences; test bias

Standardized cognitive ability testing has a far-reaching impact on the lives of individuals, various ethnic groups, and our society as a whole. Testing outcomes guide educational, career, and social paths. For example, a disproportionately high number of ethnic students are placed in special education programs (MacMillan, Gresham, & Sipersein, 1993), whereas a relatively low number are placed into gifted programs (Ford-Harris, 1993). The ongoing debate over the possible presence of cultural influences (i.e., attitudes, values, beliefs, and behaviors of a

---

This research comprised a major portion of the first author's dissertation project in partial fulfillment of a doctoral degree in clinical psychology and is registered with *Dissertation Abstracts International*. Portions of the information contained in this report were presented at the 40th Annual Meeting of the Society for Psychophysiological Research, San Diego, California, October 2000, and the 109th American Psychological Association Annual Convention, San Francisco, August 2001. This study was supported by the Minority Dissertation Research Grants in Mental Health, National Institute of Mental Health Grant, 1 R03 MH58476-01A1 to the first author. Additional

Assessment, Volume 12, No. 3, September 2005 303-319
DOI: 10.1177/1073191105276674

group of people) and the appropriate use of current eurocentrically developed assessment measures with ethnic minority and other cultural groups continues (for reviews, see Armour-Thomas, 2003; Neisser et al., 1996; Suzuki, Meller, & Ponterotto, 1996; Suzuki & Valencia, 1997; Valencia & Suzuki, 2001). Thus, the need for further research on the role of cultural influence in testing and the development of more culture-fair assessment tools persists (Suzuki & Valencia, 1997; Valencia & Suzuki, 2001).

It has been difficult to provide empirical evidence for a cultural bias in testing. Much research and discussion of bias in standardized IQ testing has addressed the difference between African American and Caucasian American IQ scores, but other ethnic groups have also shown lower IQ scores than the Caucasian American population. For example, Hispanic groups are reported to score about halfway between African American and Caucasian American groups; that is, approximately one-half standard deviation below Caucasian American groups (Suzuki & Gutkin, 1993; Whitworth & Gibbons, 1986). Prevailing arguments for group differentials on IQ tests include a true group differential, a reflection of test bias, and a pervasive societal bias that affects informal and formal education (Neisser et al., 1996). Ethnic group differences in IQ have been attributed to a genetic hypothesis, which states that lower IQ scores exhibited in one ethnic group as compared to another group are due to a genetic disposition toward lower cognitive ability levels (Jensen, 1969, 1973; Rushton, 1988). However, reviews investigating genetic contributions to group differences in cognitive ability have criticized the methodology as problematic, findings as inconsistent, and interpretations as inconclusive (Brody, 1992; Loehlin, 2000; Suzuki & Valencia, 1997). Alternatively, cognitive ability tests may be susceptible to cultural biases because they require shared values, shared knowledge, and shared communication (Greenfield, 1997).

In a recent review of the testing literature, Valencia, Suzuki, and Salinas (2001) studied 62 empirical investigations of cultural bias in cognitive ability measures. The majority of the studies (71%) stated nonbiased findings, but 29% found biased or mixed results. The authors also examined the studies' psychometric properties (reliability and validity) and noted several methodological and generalizability concerns. Many studies have investigated criterion-related validity issues in addressing ethnic group differences, but construct validity was also highlighted in the Valencia et al. (2001) review as integral to the issue of test bias because both have implications for the appropriate inferences to draw from test score outcomes. The authors concluded that cultural bias in cognitive ability assessment remains an open issue.

Cole and Moss (1989) provided a construct-validity-oriented definition of test bias that moves beyond many criterion-related definitions by incorporating differential validity for subgroups: "Bias is differential validity of a given interpretation of test scores for any definable, relevant subgroup of test takers" (p. 205). If a test were found to be significantly associated with an established measure of cognitive ability for a Caucasian American group, but not for an ethnic group, then the test would be biased. Demonstrating differential validity between groups on cognitive ability tests breaks the circular arguments in which criterion validity has been trapped (e.g., when other criterion measures are also vulnerable to cultural factors).

In addition to improvements in the construct validity of assessment instruments, the development of tests tapping basic cognitive processing efficiency may also facilitate advancements in the measurement of human intelligence (Deary & Stough, 1996; Fagan, 1992, 2000; Fagan & Haiken-Vasen, 1997). According to the cognitive/neural efficiency hypothesis, individuals with higher level intellectual abilities process information at a basic level more efficiently (i.e., more rapidly and using less mental effort) relative to individuals with lesser intellectual abilities (Davidson & Downing, 2000; Haier, Siegel, Tang, Abel, & Buchsbaum, 1992; A. E. Hendrickson, 1982; D. E. Hendrickson, 1982; Schafer, 1982). The visual backward-masking (VBM) task is one task used extensively in this research to investigate the efficiency of the early stages of human visual information processing (for reviews, see Deary & Stough, 1996; Kranzler & Jensen, 1989). The VBM task is used to quantify the amount of time required to pass information through the sensory register (inspection time [IT]). The VBM procedure consists of a rapidly presented target stimulus (e.g., letters or lines of different length), a varying length of vacant time (interstimulus intervals typically ranging from 20 milliseconds [ms] to 700 ms), and a masking stimulus that typically completely covers the spatial presence of the target stimulus (Saccuzzo, 1993). A condition in which only the target stimulus is presented (i.e., no-mask condition) is often administered to isolate the individual's ability to process the target stimulus without the effects of the masking stimulus. The individual is usually asked to decide which of a

---

support was provided in part by the Department of Defense's Multidiscipline University Research Initiative (MURI) to Dr. Sandra Marshall in collaboration with George Mason University, the National Institute of Mental Health Training Grant #T32 MH19934 to Dr. Dilip Jeste's Geropsychiatry Intervention Research Center at the Veterans Affairs San Diego Healthcare System, National Institute of Mental Health Grant, 1 R01 MH61381 to Dr. Eric Granholm, and by the Department of Veterans Affairs. Address reprint requests to Steven P. Verney, Ph.D., Department of Psychology, MSC03-2220, 1 University of New Mexico, Albuquerque, NM 87131-0001, USA; e-mail: sverney@unm.edu.

pair of stimuli were presented as targets (i.e., forced-choice paradigm). IT on the VBM task accounts for about 20% of the variance in intelligence scores (Deary & Stough, 1996; Kranzler & Jensen, 1989; Longstreth, Walsh, Alcorn, Szeszulski, & Manis, 1986; Nettlebeck, Edwards, & Vreugdenhil, 1986). Early visual information processing tasks such as IT, which concentrate on the first few stages of processing, appear to be less influenced by cultural and social learning factors than are higher order cognitive measures (Deary & Stough, 1996; Fagan, 2000; Fagan & Haiken-Vasen, 1997).

There have been only a few studies investigating information processing efficiency and ethnicity (e.g., Bosco, 1972; Fagan & Haiken-Vasen, 1997; Haiken-Vasen, 1995; Lancy, 1983; Lynn & Holmshaw, 1990; Seagrim & Lendon, 1980). A comprehensive study by Saccuzzo, Johnson, and Guertin (1994) evaluated 160 children composing equal samples of four ethnic groups on a battery of information processing tasks, including the VBM. Efficiency of processing on the VBM was found to be the strongest correlate of both IQ and membership in a gifted program. Furthermore, no ethnic differences were found for IT, and giftedness was essentially independent of ethnic background. Thus, the VBM task has shown promise as a culture-fair measure of individual differences in general cognitive capacities.

Although behavioral data measured during information processing tasks (e.g., correct response) provide information about an individual's processing output, psychophysiological measures have the potential to assess the individual's processing capacities, such as the amount of mental effort that was allocated to produce the correct response. Many lines of psychophysiological research have explored the cognitive/neural efficiency hypothesis. For example, negative correlations have been found between metabolic rates and IQ measures in positron-emission tomography (PET) studies (e.g., Haier et al., 1988, 1992; Parks, Lowenstein, & Dondrell, 1988) and between mental ability and various event-related potential measures in electroencephalographic (EEG) studies (reviewed in Vernon, Wickett, Bazana, & Stelmack, 2000). Thus, unique information about an individual's cognitive processing can be obtained by recording psychophysiological measures during cognitive task performances while simultaneously gathering more traditional behavioral measures (e.g., correct response, reaction time).

Pupillary responses are a reliable measure of processing load and processing resource allocation. In a review article of pupillometry studies of cognitive processing, Beatty (1982) established that task-evoked pupillary responses reflect within-task, between-task, and between-individual variations in processing demands. Pupillary dilation responses on tasks of working memory (e.g., digit span), language (e.g., grammatical reasoning, word match, sentence encoding), reasoning (e.g., multiplication), and perception (e.g., discrimination, detection) have been shown to reflect processing demands. Ahern and Beatty (1979) demonstrated that individuals with higher SAT scores had smaller pupillary dilations than individuals with lower SAT scores during multiplication tasks. They concluded that individuals with more intellectual abilities appear to process information with more efficiency. More recently, Granholm, Asarnow, Sarkin, and Dykes (1996) used a digit recall task to demonstrate that as processing load (i.e., digit span) increased, pupil size recorded during the task also increased until it reached an asymptote at resource limits (memory capacity of 7 ± 2 items), and then declined with overload. Thus, pupillary responses appear to provide an index of processing capacity limitations.

In a previous study, we demonstrated that pupillary dilation response measured processing resource allocation on the VBM task (Verney, Granholm, & Dionisio, 2001). Pupillary dilation responses were significantly greater during task performance (cognitive load) than during a passive stimulus viewing condition (no load). Consistent with models of early visual information processing (Michaels & Turvey, 1979; Phillips, 1974), pupillary dilation responses were significantly greater in the 317 ms condition than the no-mask condition, suggesting that the irrelevant masking stimulus demanded extra processing resources when it followed the target by more than 100 ms.

The psychological assessment field needs new ideas and methodologies that take advantage of new technologies in order to progress toward culture-fair cognitive ability assessment. Promising new methodologies to accomplish this include information processing (VBM detection accuracy scores) and psychophysiological measures (pupillary response measures) of cognitive ability. By relating efficiency of information processing to higher order processing (i.e., standardized IQ test) within two different cultural groups (Mexican American and Caucasian American), we can better understand the relationship between culture and cognitive capacity. This study proposes to use a novel method of using information processing and psychophysiological approaches to investigate the clinically significant issue of evaluating equivalence of construct validity in cognitive ability measures.

In this study it was hypothesized that after controlling for socioeconomic status (SES), higher Wechsler Adult Intelligence Scale–Revised (WAIS-R) Full Scale IQ (FSIQ) scores would be associated with greater detection accuracy on the VBM task in the total sample. SES was used as a covariate because it is correlated with cognitive ability and has been shown to moderate ethnic difference effects (Valencia & Suzuki, 2001). If confirmed, this would replicate previous findings that behavioral measures of information processing efficiency are strongly related to cognitive ability (Deary & Stough, 1996; Kranzler & Jensen, 1989; Saccuzzo et al., 1994). It was also hypothesized that pupillary dilations elicited by the VBM task masking stimulus at longer stimulus-onset asynchrony conditions would significantly add to the prediction in FSIQ scores above and beyond that provided by detection accuracy, with higher FSIQ scores associated with smaller pupillary responses. If confirmed, this would replicate previous findings that individuals with greater cognitive abilities are able to perform cognitive tasks with less effort (Ahern & Beatty, 1979). This would also provide further evidence that individuals who wastefully allocate resources to irrelevant information perform more poorly on IQ tests (e.g., Cha & Merrill, 1994; McCall, 1994; Merrill & Taube, 1996). Finally, it was hypothesized that these relationships between FSIQ and detection accuracy and pupillary response measures on the VBM task would be stronger for the Caucasian American student sample than for the Mexican American student sample. This hypothesis is consistent with the definition and experimental approach of Cole and Moss (1989). If confirmed, this would indicate that WAIS-R FSIQ scores are a less valid measure of cognitive ability for Mexican American students than for Caucasian American students.

METHOD

Participants

Undergraduate male and female students (n = 101) were recruited from introductory psychology courses at San Diego State University (SDSU). Participants were offered class credit and monetary compensation for their time and efforts and provided informed consent. The human subject committees of the University of California, San Diego (UCSD), and SDSU approved this study.

Gender and ethnic inclusion. Two different ethnic populations, Caucasian American and Mexican American undergraduate students, were recruited for the study, because these populations represent the largest proportions in the SDSU community. Students who stated availability for research studies at the university were contacted if they had self-identified as belonging to either of these two groups. The Scale of Ethnic Experience (SEE) (Malcarne, Chavira, & Fernandez, 2004), which asks participants to identify their ethnic group membership, was used to assess their ethnic group status at the time of testing. The term "Caucasian American" will be used to refer to those individuals who identified themselves as non-Hispanic White and European American.

Participant exclusion criteria. Five individuals were dropped from the study due to medical or physical reasons: history of loss of consciousness over 2 minutes (n = 1), attention deficit disorder (n = 1), epileptic seizures (n = 1), learning disabilities (n = 1), and falling asleep during testing (n = 1). One Peruvian student, who did not meet criteria for inclusion into the Mexican American group, was dropped from the study. Twelve participants who presented for the study were excluded from all the analyses due to excessive eye blink artifacts (analyzable trials less than 40%; n = 2), technical difficulties with the eye-tracking instrument (n = 2), abnormal tonic pupil measurements (resting diameter outliers greater than 3 standard deviations from the mean of all respondents; n = 2), or extremely poor VBM performance (at-chance no-mask condition or no-mask performance outliers greater than 3 standard deviations from the mean of all participants; n = 6). Eight participants were excluded from the all the analyses due to regression outlier criteria (see below). All participants demonstrated at least 20/30 visual acuity (corrected or noncorrected) as assessed by a Snellen wall chart. No participant reported a history of substance abuse, or smoking cigarettes or drinking caffeinated beverages within 2 hours prior to the testing session.

Table 1 presents means and standard deviations and t test comparisons between the two groups on relevant demographic variables. The two groups did not differ significantly in age, education, or gender (see Table 1).

Apparatus

A head-mounted infrared corneal-reflection-pupil-center eye-tracking system, the Applied Science Laboratories Model 4000SU HMO, was used to gather pupillometric data from the left eye during the VBM task performance. Pupil diameter was sampled at 60 hertz (Hz) (approximately every 16.7 ms). A 17-inch Super Video Graphics Adapter (SVGA) monitor controlled by a personal computer (PC) was used to administer the VBM task, and the keyboard arrow keys provided the right and left response keys.

Procedure

VBM task. The VBM task was administered on a PC-based system in order to accommodate the eye-tracking system setup. The constraints of the computer monitor refresh rate (i.e., 60 Hz; 16.7 ms) prevented the attainment of the traditional tachistoscopic-administered measurement of IT on the VBM task. The IT procedure consists of the manipulation of the critical stimulus duration (usually in ms increments) in order to obtain an individual's criterion level of accuracy (e.g., 80%; for a review of the IT litera-

TABLE 1
Variable Comparisons (Means and Standard Deviations) for the Total and Full Scale IQ (FSIQ)-Compatible Samples

	Total Sample					*FSIQ-Compatible Samples*[a]				
Variable	*Mexican American* (n = *35*)		*Caucasian American* (n = *40*)		*Statistic*	*Mexican American* (n = *21*)		*Caucasian American* (n = *35*)		*Statistic*
Age (years)	18.9	(1.7)	18.3	(0.8)	$t(73) = 1.79$	18.8	(1.8)	18.3	(0.9)	$t(54) = 1.11$
Education (years)	12.5	(0.8)	12.5	(0.7)	$t(73) = 0.08$	12.5	(0.7)	12.3	(0.6)	$t(54) = 0.85$
Gender (% female)	56.4		52.3		$\chi^2(1) = 0.88$	52.3		57.1		$\chi^2(1) = 0.12$
Family income[b]	35.1	(18.9)	57.5	(7.9)	$\chi^2(5) = 29.6$**	38.1	(17.4)	57.1	(8.4)	$\chi^2(4) = 20.1$**
Father's education (years)	10.7	(4.0)	14.5	(2.2)	$\chi^2(5) = 24.7$**	11.9	(3.5)	14.5	(2.4)	$\chi^2(6) = 13.3$*
Mother's education (years)	11.5	(4.0)	14.0	(2.1)	$\chi^2(5) = 18.0$**	11.8	(3.4)	14.3	(1.6)	$\chi^2(6) = 21.5$**
Socioeconomic status	3.7	(1.4)	5.4	(0.6)	$t(73) = 6.24$**	4.1	(1.2)	5.4	(0.6)	$t(54) = 4.86$**
Ethnic identification[c]	3.51	(0.68)	3.09	(0.40)	$t(73) = 3.34$**	3.63	(0.66)	3.12	(0.41)	$t(54) = 3.48$**
Perceived discrimination[c]	3.05	(0.42)	2.27	(0.34)	$t(73) = 8.91$**	3.11	(0.50)	2.29	(0.34)	$t(54) = 7.37$**
Mainstream comfort[c]	3.51	(0.77)	4.19	(0.53)	$t(73) = 4.49$**	3.60	(0.84)	4.14	(0.52)	$t(54) = 2.98$**
Social affiliation[c]	2.44	(0.76)	2.39	(0.56)	$t(73) = 0.32$	2.51	(0.75)	2.43	(0.59)	$t(54) = 0.47$
English as 1st language (%)	62.9		100.0		$\chi^2(1) = 18.0$**	67.0		100.0		$\chi^2(1) = 13.3$**
Bilingual (%)	65.7		2.5		$\chi^2(1) = 34.2$**	57.1		2.9		$\chi^2(1) = 21.7$**
Generation in U.S.	2.80	(1.21)	4.45	(0.90)	$\chi^2(4) = 32.0$**	3.05	(1.24)	4.54	(0.74)	$\chi^2(4) = 21.2$**
WAIS-R Full Scale IQ	99.0	(10.0)	105.6	(6.9)	$t(73) = 3.30$**	103.2	(5.6)	104.0	(5.6)	$t(54) = 0.51$
WAIS-R Verbal IQ	97.8	(8.6)	105.1	(10.4)	$t(73) = 3.31$**	102.0	(5.6)	102.9	(8.0)	$t(54) = 0.49$
WAIS-R Performance IQ	100.5	(13.2)	106.0	(10.2)	$t(73) = 2.02$*	103.6	(9.4)	104.4	(8.9)	$t(54) = 0.32$
VBM detection accuracy[d]	82.1	(11.7)	86.0	(8.2)	$t(73) = 1.65$	86.3	(7.1)	85.6	(8.6)	$t(54) = 0.30$
VBM mask pupil response[e]	0.27	(0.80)	0.09	(0.55)	$t(73) = 1.18$	0.05	(0.57)	0.12	(0.58)	$t(54) = 0.50$

NOTE: WAIS-R = Wechsler Adult Intelligence Scale–Revised; VBM = visual backward masking. Standard deviation in parentheses.

a. FSIQ-compatible samples are a subset of the total sample in which the two groups were made comparable on WAIS-R FSIQ.

b. In thousands of dollars.

c. Scale of Ethnic Experience factor.

d. Average VBM detection accuracy (% correct).

e. Average VBM mask pupillary response (principal components analysis factor score).

*$p < .05$. **$p < .01$.

ture, see Deary & Stough, 1996). Thus, detection accuracy during a variety of stimulus onset asynchronies (SOA) on the VBM task was used in this study as an index of information processing efficiency. Participants were asked to identify which of two target lines was longer (i.e., forced-choice paradigm). Target and masking stimuli consisted of black lines on a white background to reduce screen glare effects and minimize the change that could be associated with the pupil light reflex. The target stimulus consisted of two adjacent vertical lines presented in the center of the computer screen, 1.7 centimeters (cm) apart. For every trial, one of the two lines (right or left) was longer than the other (2.7 vs. 2 cm) and vertically offset in height in one of six different target configurations. The upper end point of the "short" line could be higher, equal, or lower than the upper end point of the "long" line, and the short line could be either the right or left line. Also, only one end point (upper or lower) of one target line could be in alignment with the same end points of the masking lines. In this manner, the effects of apparent movement or flicker as a mask-breaking cue may be reduced (Alexander & Mackenzie, 1992). The long and short lines were randomly blocked in series of 12 trials (so that each of the six offset configurations was presented twice in every sequence of 12 trials). The masking stimulus was composed of two 4-cm-long, parallel lines that completely spatially replaced the target stimulus lines with SOAs of either 50, 67, 100, 134, 317, or 717 ms. A condition consisting of only the target stimulus with no masking stimulus (no-mask condition) was included as a control condition. These SOAs, composing a typical range in the backward-masking literature and covering the range of fundamental visual backward-masking mechanisms (Breitmeyer & Ganz, 1976; Michaels & Turvey, 1979; Phillips, 1974), were bounded by the 60 Hz refresh rate of the monitor and were timed to display in accordance with the top of the refresh cycle. Twenty trials were administered for each condition, resulting in 140 test trials. The six SOA conditions along with the no-mask condition were presented in four blocks, with each block containing a set of 5 trials of each condition. The conditions were randomized within each block, resulting in the following counterbalanced sequence: 134, no mask, 67, 100, 317, 717, 50, 134, 717, 317, 50, no mask, 100, 67, 50, no mask, 717, 100, 317, 67, 134, 317, 50, 67, 717, 134, no mask, 100 ms. Thus, presentation confounds (e.g., performance fatigue, pupillary response habituation) between conditions during the task were diminished. Both target and mask were of equal duration (16.7 ms; one 60 Hz screen refresh rate).

A calibration was first conducted to ensure participant-pupillometer agreement on the center of visual field. At the beginning of each trial, a blue fixation square (0.85 cm × 0.85 cm) was presented in the center of the monitor (with a white screen background) for 1 second (sec) along with a high-pitched tone (1,500 Hz for 500 ms). The fixation square and tone served as visual and auditory cues to warn the participant to prepare for the trial's target stimulus. Instructions were given to press either the right or left arrow keys on the keyboard to indicate which of the two test lines was longer. Both detection accuracy and speed was emphasized with the instruction, "Try to be as accurate as you can, but also be as fast as you can." Three seconds after the onset of the target stimulus, a low-pitched tone (800 Hz for 500 ms) functioned as an auditory cue signaling the end of the trial. The intertrial interval was set at 3 sec. Participants were asked to refrain from blinking during the trial period marked by the two auditory signals (i.e., high and low beeps).

Prior to the test portion of the task, the participant was given 21 practice trials. The practice trials began with the easiest SOA conditions; namely, a no-mask trial followed by a 717 ms and a 317 ms SOA trial. The SOA durations of the remaining 18 trials were randomly blocked. The first 12 practice trials provided computer-automated feedback regarding the correctness of participant's response. Feedback was not provided during the test phase of the study. A moment of rest (i.e., approximately 10 to 15 sec) was allowed after each presentation of 10 trials, and each participant was allowed a few minutes to rest halfway through the test. The entire task (i.e., instructions, practice, and test) took typically less than 35 minutes, with the test portion taking about 22 minutes.

WAIS-R performance. The Satz-Mogel Abbreviated Procedure (Satz & Mogel, 1962) for the WAIS-R (Wechsler, 1981) was administered in English to each participant. The WAIS-R was used in this study because much of the ethnic-difference literature had been conducted using the WAIS-R in the 1960s through the 1980s (see Valencia, Suzuki, & Salinas, 2001). Scores derived using the Satz-Mogel Abbreviated Procedure have demonstrated high short form-complete form correlations with the complete WAIS-R FSIQ (.95), Verbal IQ (VIQ) (.94), and Performance IQ (PIQ) (.89) (Silverstein, 1982). In this short version, all subtests are given, but with only a portion of the questions administered (e.g., every third or every other item), thus requiring half the administration time.

The Wide Range Achievement Test, 3rd edition (WRAT3). The WRAT3 (Wilkinson, 1993) Reading subtest was administered to ensure that each participant met the minimum English reading ability level for valid WAIS-R results (Wechsler, 1981). All participants met the English reading level minimum requirement. Furthermore, there was no significant difference between the Mexican American sample and the Caucasian American sample on the WRAT3 Reading subtest: Mexican American sample *M* = 47.83, *SD* = 4.06; Caucasian American sample *M* = 47.90, *SD* = 3.61; *t*(73) = .08, *ns*. WRAT3 Reading subtest scores were significantly correlated with WAIS-R FSIQ scores in the Caucasian American sample, *r*(40) = .31, $p < .05$, but not in the Mexican American sample, *r*(35) = .26, *ns*. These correlations were not significantly different between the two groups, $z_r = .22$, *ns*.

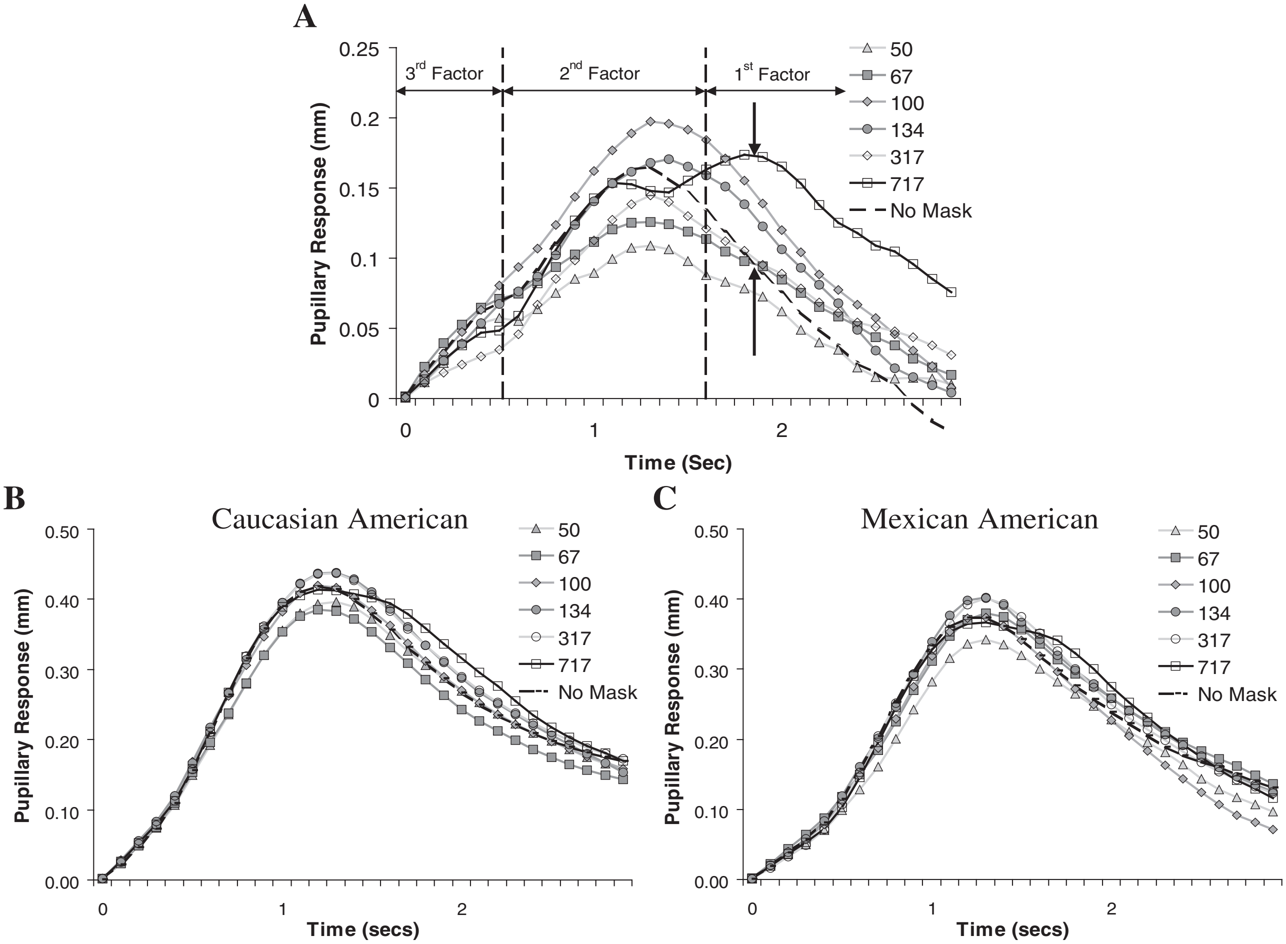
SES. SES is one of the major variables accounting for intellectual differences within and between ethnic groups, often transcending the predictor of ethnicity (Suzuki & Valencia, 1997). SES was estimated by averaging the self-reported variables of the participant's mother and father's education level and family income. A relative score for each participant was calculated by averaging the index of each categorical variable. For example, if a participant had a mother with 14 years of education, corresponding to an index of 5 (some college), a father with 11 years of education (index = 3, some high school beyond Grade 8), and a family income of $45,000 (index = 4, $40,000 to $50,000), his or her resulting SES score would be 4.0.

SEE. The SEE (Malcarne et al., 2004) is a 32-self-report-item questionnaire that can be used across ethnic and racial groups. The SEE utilizes a 5-point Likert-type scale to tap into a range of cultural and mainstream experiences. Factor analyses have yielded four stable factors of the SEE: Ethnic Identity, Perceived Discrimination, Social Affiliation, and Mainstream Comfort. The SEE has demonstrated good to excellent internal consistency across various ethnic groups, satisfactory test-retest reliability, and strong concurrent validity with other relevant ethnic and acculturation measures.

Analyses

VBM pupillary response. Graphic displays of raw pupil diameter data were first visually inspected for gross artifacts by a trained technician. Fewer than 4.3% of the test trials were discarded due to major artifacts (e.g., technical difficulties in data acquisition, obstruction of the eye image) or excessive eye blinking. A computer algorithm was used to remove eye blinks and other minor artifacts from other trials by linear interpolation (Beatty & Lucero-Wagoner, 2000). A 7-point smoothing filter was then passed over the data. For each participant, an average pupillary response was calculated for the artifact-free trials of each SOA condition. To fully and objectively examine the pupillary response waveform across the 3-second trial, and to eliminate the effects of individual differences in resting pupil size and pupil mobility, a varimax rotation principle components analysis (PCA) was performed on 180 time points (i.e., 3 sec) of the pupil response waveform time locked to stimulus onset across the seven masked and no-mask conditions for all participants. PCA is often used in psychophysiological data as a method of reducing the large number of data time points to a small number of meaningful factors. A varimax rotation, one of the most commonly used rotation methods in PCA analysis, minimizes the complexity of factors by maximizing variance of loadings on each factor. Orthogonal components are extracted that summarize the patterns of correlations in the observed data points. Resulting factor scores are the correlations between the data time points and the PCA factors. A successful PCA will account for a large amount of variance in the data and yield interpretable factors that make sense (Tabachnick & Fidell, 1996). Three prominent stable factors emerged (see Figure 1A) accounting for 95.3% of the variance in the pupillary response data. As indicated by the squared multiple correlations, all factors were internally consistent and well defined by the data (the lowest of the squared multiple correlations for factors from data was .67). The three components formed a linear time course of the pupillary response waveform (see Figure 1A): (a) an early component from 0 to 0.7 sec (the third rotated factor; eigenvalue, or percentage of variance explained, of 5.8); (b) a middle component from 0.7 to 1.53 sec (second rotated factor; eigenvalue = 20.8); and (c) a late component from 1.53 to 3.0 seconds (first rotated factor; eigenvalue = 68.7). The third rotated factor occurs at a time frame when a response to an illumination change typically occurs (Loewenfeld, 1999). The second rotated factor occurs at the time window typically taken to represent resource allocation to the task performance (e.g., discriminating and evaluating the target lines, making a decision as to which line is longer, and generating a motor response) (Beatty, 1982; Steinhauer & Hakerem, 1992). The first rotated factor, which included the pupillary response's return to baseline, likely indexed resource allocation to the mask in the longer SOA conditions (i.e., greater than 100 ms) (Verney et al., 2001). Based on our previous findings (Verney et al., 2001), we were most interested in the first rotated factor because this factor allowed us to isolate the resource allocation devoted to the masking stimulus, an irrelevant stimulus to the task. Wastefully allocating resources to irrelevant information has been associated with lower cognitive ability performance (e.g., Cha & Merrill, 1994; McCall, 1994; Merrill & Taube, 1996). The masking stimulus becomes a distinct percept from the target stimulus at SOA conditions greater than 100 ms (Michaels & Turvey, 1979; Phillips, 1974). Therefore, the difference between the average of the longer SOA conditions (i.e., 317 and 717 ms conditions) and the no-mask condition in the first rotated factor was used as an index of the amount of resource allocation devoted to the masking stimulus (average VBM mask pupillary response).

FIGURE 1
Raw Pupillary Responses (in millimeters [mm]) From Baseline for the Six Stimulus Onset Asynchronies (SOA) Conditions and the No-Mask Condition Elicited During the Visual Backward-Masking (VBM) Task



NOTE: (A) The three principal components analysis (PCA) rotated factors for one respondent. The difference between the average of the longer SOA conditions (i.e., 317 and 717 milliseconds [ms] conditions) and the no-mask condition for the first factor was used as an index of the amount of resource allocation devoted to the masking stimulus (average VBM mask pupillary response). The arrow indicates resources devoted to the mask in the longest SOA condition. Averaged raw pupillary responses for (B) the Caucasian American group ($n = 35$) and (C) for the Mexican American group ($n = 21$).

VBM detection accuracy. The number of correct responses on the VBM task was recorded for each of the six masked SOA conditions and the no-mask condition. To maximize the sensitivity of overall detection accuracy to individual differences in VBM performance, conditions with floor and ceiling effects were discarded. Floor effects were determined by calculating the probability that performance in an SOA condition differed significantly from chance-level performance (50%) using 95% confidence intervals. Both the 50 and 67 ms SOA conditions were at chance-level detection accuracy and were discarded when calculating overall detection accuracy. Performance in the no-mask condition did not differ significantly from perfect accuracy and had low variability ($M = 94.3$, $SD = 6.3$). Thus, the no-mask condition was at ceiling and discarded when calculating overall detection accuracy. Average VBM detection accuracy was, therefore, defined as the average percentage correct for the intermediate and longer masked SOA conditions (i.e., 100, 134, 317, and 717 ms conditions).

Statistical analysis. All regression residuals met criteria for normality, linearity, and homoscedasticity. Univariate outlying scores (scores > 3 *SD*) were assigned a score that was one unit larger (or smaller) than the next most extreme score in the distribution (Tabachnick & Fidell, 1996). This procedure only resulted in adjusting one Caucasian American participant's WAIS-R FSIQ score downward. Bivariate and multivariate outliers were identified through studentized residuals (studentized residual > 2) and were dropped. Four Caucasian American and 4 Mexican American participants were dropped as multivariate outliers.

The dependent variable was WAIS-R FSIQ. Four predictor variables were calculated for each participant: (a) SES was defined as the average ratings of family income, father's education level, and mother's education level, (b) average VBM detection accuracy, (c) average VBM mask pupillary response, (d) ethnicity. Neither gender nor VBM median response time significantly influenced the analyses below; thus, these variables are not included in the analyses.

Several multiple linear regressions were used to examine relationships among various combinations of the variables in different subgroups. The total sample was used to test the contributions of the information processing efficiency variables to FSIQ. To compare information processing efficiency within the two groups, two subgroups of the total sample, which were compatible on FSIQ as described below, were used. To control for Type I error, significance levels for the regressions on group comparisons were Bonferroni corrected to $\alpha = .025$. Repeated measures analysis of variance (ANOVA) was used to analyze the predictor variables and were adjusted for violations of sphericity using the Huynh-Feldt adjustment where appropriate.

RESULTS

Ethnic Group Comparisons

WAIS-R FSIQ, VIQ, and PIQ. Table 1 presents the WAIS-R FSIQ, VIQ, and PIQ scores for the Caucasian American and Mexican American total and FSIQ-compatible samples. Caucasian American participants performed significantly higher than Mexican American participants on FSIQ, VIQ, and PIQ in the total samples. When the two groups were of comparable FSIQ (i.e., FSIQ-compatible samples, as outlined below), there were no significant differences on FSIQ, VIQ, or PIQ.

SES. Table 1 presents the SES variable and its components (i.e., mother's education level, father's education level, family income) for the total and FSIQ-compatible samples of Caucasian American and Mexican American participants. Caucasian American participants had significantly higher SES for both the total and the FSIQ-compatible samples than did the Mexican American participants.

SEE. Table 1 presents the SEE (Malcarne et al., 2004) factor scores of Ethnic Identity, Perceived Discrimination, Mainstream Comfort, and Social Affiliation as well as the self-reported items of having learned English as the first language, being bilingual, and the number of generations the participant's family has resided in the United States (Generations in US) for the total and FSIQ-compatible samples of Caucasian American and Mexican American participants. For both the total and FSIQ-compatible samples, Mexican American participants reported being significantly more ethnically identified, perceiving more discrimination, being less comfortable in mainstream U.S. society, being less likely to have learned English as a first language, being more likely to be bilingual, and having had fewer generations lived in the United States than their Caucasian American counterparts.

VBM detection accuracy. Table 1 presents the average VBM detection accuracy scores, which did not differ significantly between the Caucasian American and Mexican American participants for the total and FSIQ-compatible samples. Table 2 presents VBM detection accuracy scores for the six SOA masking conditions and the no-mask condition for the total and FSIQ-compatible samples. All samples show a typical profile for the VBM task, with at-chance performance in the shorter SOAs and higher percentage correct responses for the longer SOAs and the no-mask conditions. A 2 (Ethnicity) × 7 (SOA) repeated-measures ANOVA conducted on the percentage of correct responses in the total samples resulted in a significant main effect for SOA, $F(6, 367.7) = 144.02$, $p < .01$, $\eta^2 = .66$, but the main effect for ethnicity was not significant; $F(1, 73) = 2.82$, *ns*, $\eta^2 = .04$. The SOA × Ethnicity interaction was also not significant, $F(6, 367.7) = 1.60$, *ns*, $\eta^2 = .02$. A similar pattern of ANOVA results was found for the FSIQ-compatible samples: SOA, $F(6, 251.6) = 129.53$, $p < .01$, $\eta^2 = .71$; ethnicity, $F(1, 54) = .00$, *ns*, $\eta^2 = .00$; SOA × Ethnicity, $F(6, 251.6) = 1.01$, *ns*, $\eta^2 = .02$.

VBM mask pupillary response. Figure 1 presents the averaged raw pupillary responses for the Caucasian American group (see Figure 1B) and the Mexican American Group (see Figure 1C). Table 1 presents the average VBM mask pupillary response, which did not differ significantly between the Caucasian American and Mexican American participants in either the total or the FSIQ-compatible samples. Table 2 presents pupillary responses for the late mask component (first rotated PCA factor) for the six SOA masking conditions and the no-mask condition for the total samples. A 2 (Ethnicity) × 7 (SOA) repeated measures ANOVA conducted on these data resulted in a significant main effect for SOA, $F(6, 398) = 3.56$, $p < .01$, $\eta^2 = .05$. The ethnicity main effect was not significant, $F(1, 73) = .17$, *ns*, $\eta^2 = .00$, nor was the SOA × Ethnicity interaction,

TABLE 2
Visual Backward Masking (VBM) Task Detection Accuracy and Pupillary Response Means and Standard Deviations for the Stimulus-Onset Asynchrony (SOA) Conditions in the Total and Full Scale IQ (FSIQ)-Compatible Caucasian American and Mexican American Samples

	Stimulus-Onset Asynchrony						
	50 ms	*67 ms*	*100 ms*	*134 ms*	*317 ms*	*717 ms*	*No Mask*
Total sample							
VBM detection accuracy (% correct)							
Caucasian American (*n* = 40)	64.7 (13.3)	63.6 (14.2)	74.0 (14.5)	81.7 (12.4)	94.2 (8.2)	94.1 (7.6)	94.6 (6.5)
Mexican American (*n* = 35)	58.7 (13.8)	63.9 (12.4)	73.9 (15.6)	76.1 (12.9)	88.5 (14.0)	90.0 (13.8)	94.0 (6.0)
VBM pupillary response (principal components analysis factor score)							
Caucasian American (*n* = 40)	0.02 (1.04)	0.01 (1.19)	0.04 (1.15)	0.09 (1.10)	0.07 (1.11)	0.18 (0.98)	0.04 (0.99)
Mexican American (*n* = 35)	–0.07 (0.99)	–0.02 (0.98)	–0.27 (0.85)	0.04 (0.85)	0.01 (0.93)	0.26 (0.94)	–0.14 (0.96)
FSIQ-compatible samples							
VBM detection accuracy (% correct)							
Caucasian American (*n* = 35)	63.9 (12.8)	64.0 (12.2)	73.3 (14.8)	81.1 (12.7)	94.3 (8.5)	93.9 (8.0)	94.6 (6.8)
Mexican American (*n* = 21)	59.3 (15.3)	64.7 (13.6)	77.1 (13.0)	79.9 (10.6)	93.9 (8.5)	94.3 (7.1)	94.9 (4.8)
VBM pupillary response (principal components analysis factor score)							
Caucasian American (*n* = 35)	0.07 (1.0)	0.07 (1.21)	0.01 (1.15)	0.05 (1.12)	0.09 (1.16)	0.19 (1.01)	0.02 (1.02)
Mexican American (*n* = 21)	–0.08 (0.86)	0.07 (1.04)	–0.27 (0.83)	–0.02 (0.80)	–0.07 (0.79)	0.03 (0.72)	–0.07 (0.91)

NOTE: ms = milliseconds. Standard deviation in parentheses.

$F(6, 398) = 1.22$, *ns*, $\eta^2 = .02$. Consistent with our previous study (Verney et al., 2001) follow-up analyses (Bonferroni's method) found that only the latest SOA condition (717 ms) was significantly greater than the no-mask condition, $t(74) = 2.84$, $p < .01$, suggesting that the 717 ms condition required greater resource allocation over that of the no-mask condition. A 2 (Ethnicity) × 7 (SOA) repeated measures ANOVA conducted on these data for the FSIQ-compatible samples did not result in significant main or interaction effects; SOA, $F(6, 324) = 1.31$, ns, $\eta^2 = .02$; ethnicity, $F(1, 49) = .18$, *ns*, $\eta^2 = .00$; SOA × Ethnicity, $F(6, 324) = 1.08$, *ns*, $\eta^2 = .02$.

Regression Analyses

VBM detection accuracy and FSIQ. A regression analysis was used to investigate the study's first hypothesis, that is, the relationship between the behavioral index of information processing efficiency, VBM detection accuracy, and general cognitive ability, WAIS-R FSIQ scores. The hierarchical regression model for the prediction of WAIS-R FSIQ in the total sample is presented in Table 3. SES was used as a covariate to better investigate the relationships of interest. SES significantly predicted FSIQ in Step 1 of the regression, accounting for 13.7% of the variance. Consistent with the study hypothesis, the addition of average VBM detection accuracy resulted in a significant model and in a significant change in R^2. Higher average VBM detection accuracy significantly predicted higher FSIQ and accounted for 17% of the variance in FSIQ beyond that accounted for by SES.

VBM pupillary response and FSIQ. The hierarchical regression was expanded to investigate the study's second hypothesis, that is, the contribution of the psychophysiological index of information processing efficiency, VBM mask pupillary response. As predicted, Step 3, adding average VBM mask pupillary response, added another 10.6% of the variance in predicting FSIQ. Lower average VBM mask pupillary response predicted higher FSIQ.

Ethnicity and FSIQ. To investigate the relationship between ethnicity and FSIQ after SES and information processing efficiency had been taken into account, ethnicity was added to the hierarchical regression in Step 4. The addition of ethnicity resulted in a significant model and a significant change in R^2 over Step 3, accounting for another 3.9% of the variance in predicting FSIQ. This suggests that predictions of FSIQ depended on group membership. This relationship is investigated further in the analyses below.

Table 3 also presents the distribution of the unique variance, sr^2, attributed to each predictor variable in the full model regression predicting FSIQ. In the presence of all predictor variables, SES did not significantly add to the FSIQ full model regression. Average VBM detection accuracy significantly accounted for 8.0% of FSIQ, and average VBM mask pupillary response significantly accounted for 11.1% of FSIQ. Therefore, average VBM detection accuracy and average VBM mask pupillary response, measures of information processing efficiency, were independently associated with overall cognitive ability. Furthermore, both information-processing variables appeared to be tapping into an overall cognitive ability level rather than a specific cognitive domain, because each variable was significantly correlated with each WAIS-R factor, VIQ, and PIQ, and no differences were found between the correlations (no z_r was greater than 1.46). VBM detection accuracy was significantly correlated with WAIS-R VIQ, $r(75) = .43$, $p < .01$, and PIQ, $r(75) = .39$, $p < .01$. VBM mask pupillary response was significantly correlated with WAIS-R VIQ, $r(75) = -.51$, $p < .01$, and PIQ, $r(75) = -.28$, $p < .05$. Ethnicity also significantly accounted for 3.9% of FSIQ. To examine this further, we investigated the contribution of information processing efficiency in FSIQ-compatible Caucasian American and Mexican American groups.

Predicting FSIQ in FSIQ-compatible samples. Regression analyses were used to investigate the differential validity hypothesis in cognitive assessment, the third stated hypothesis of this study, by analyzing the relationship between the information processing efficiency variables (VBM performance and pupillary response) and general cognitive ability (WAIS-R) for the two ethnic groups. Consistent with previous research (Suzuki & Gutkin, 1993; Whitworth & Gibbons, 1986), FSIQ distributions for the two ethnic groups (see Table 1) indicated overlapping but offset distributions with different variances. This is problematic for ethnic group comparisons because it does not allow for a direct analysis of the construct validity in question. For example, the relationship between information processing efficiency and IQ may not be a linear one. That is, individuals exhibiting lower IQ levels may be more dependent on the efficiency of information processing to perform a given task than individuals performing at higher IQ levels. Studies have suggested that the IT-IQ relationship is stronger with more cognitively impaired and older individuals (Deary, Hunger, Langan, & Goodwin, 1991; Nettelbeck & Rabbitt, 1992). Conversely, individuals performing at higher IQ levels may be processing at efficiency levels that are more task dependent than efficiency dependent. Therefore, regression analyses predicting FSIQ by information processing and pupillary response variables were conducted with a FSIQ range common to both groups. The common range was defined as ± 1 standard deviation from the total sample mean (93 to 112 FSIQ points). Thus, subgroups taken from the total

TABLE 3
Hierarchical Regressions Predicting Full Scale IQ (FSIQ) for the Total Sample and the FSIQ-Compatible Caucasian American and Mexican American Samples

		Full Model Regression Statistics					*Hierarchical Regression Statistics*			
Model	*Variable*	β	t *Value of* β	*Semipartial,* sr^2	F *Value of* sr^2	R^2	F *Value of* R^2	ΔR^2	F *Value of* ΔR^2	*Adjusted* R^2
Total sample ($n = 75$)										
Step 1	SES	.01	0.17	.000	.03	.137**	$F(1, 73) = 11.61$			.125
Step 2	Avg. Det. Ac.	.32**	3.20	.080**	10.23	.307**	$F(2, 72) = 13.66$	.170**	$F(1, 72) = 17.61$	.288
Step 3	Avg. MPR	–.37**	3.76	.111**	14.16	.413**	$F(3, 70) = 15.56$	.106**	$F(1, 71) = 12.86$	.388
Step 4	Ethnicity	.25*	2.23	.039*	4.96	.452**	$F(4, 69) = 13.50$	.039*	$F(1, 70) = 4.96$	.421
Caucasian American FSIQ-compatible sample ($n = 35$)										
Step 1	SES	–.17	0.90	.026	1.11	.001	$F(1, 33) = .02$			–.030
Step 2	Avg. Det. Ac.	–.33*	2.12	.106*	4.47	.079	$F(2, 32) = 1.36$	.078	$F(1, 32) = 2.71$	.021
Step 3	Avg. MPR	–.44**	2.83	.189**	8.01	.268*	$F(3, 31) = 3.78$	.189**	$F(1, 31) = 8.02$	.197
Mexican American FSIQ-compatible sample ($n = 21$)										
Step 1	SES	–.21	0.83	.039	.74	.012	$F(1, 19) = 0.23$			–.040
Step 2	Avg. Det. Ac.	.12	0.53	.012	.23	.021	$F(2, 18) = 0.19$	.009	$F(1, 18) = 0.17$	–.088
Step 3	Avg. MPR	–.33	1.33	.094	1.81	.115	$F(3, 17) = 0.76$	.094	$F(1, 17) = 1.80$	–.041

NOTE: SES = socioeconomic status; Avg. Det. Ac. = average visual backward masking (VBM) detection accuracy; Avg. MPR = average VBM mask pupillary response.
*$p < .05$. **$p < .01$.

samples falling in this common range formed the FSIQ-compatible groups. Table 1 presents the pertinent demographic and regression variables for the resulting FSIQ-compatible groups. The two groups did not differ significantly from each other on FSIQ, nor did they differ significantly on the predictor variables of average VBM detection accuracy or average VBM mask pupil response. However, the compatible Caucasian American participants had a significantly higher SES background than did their Mexican American counterparts.

To investigate the hypothesis that efficiency of information processing would significantly predict FSIQ scores in the Caucasian American sample, a hierarchical regression model for the FSIQ-compatible Caucasian American participants was conducted and is presented in Table 3. In Step 1, SES was not significant. Step 2 added average VBM detection accuracy and resulted in neither a significant model nor a significant change in R^2. Step 3 added average VBM mask pupillary response and resulted in both a significant model and a significant change in R^2. Information processing efficiency, as indexed by both the average VBM detection accuracy and average mask pupillary response variables, accounted for 26.7% of the variance in FSIQ above that of SES in the FSIQ-compatible Caucasian American group. The full model regression revealed that average VBM mask pupillary response significantly and independently accounted for 18.9% of the variance in FSIQ and average VBM detection accuracy significantly and independently accounted for 7.8%.

Table 3 also presents the hierarchical regression model for the FSIQ-compatible Mexican American participants. In Step 1, SES was not significant. Despite significant group differences, SES appeared not to have played a significant role in the regression analyses investigating ethnic differences. Step 2 added average VBM detection accuracy and Step 3 added average VBM mask pupillary response, but neither steps resulted in significant models or in significant changes in R^2. Consistent with the study hypothesis, significant changes in R^2 attributed to information processing efficiency in predicting FSIQ were found for the Caucasian American sample, whereas the regression model for the Mexican American sample resulted in a much weaker prediction (two fifths the effect size, R^2, for the Caucasian American sample), which was not statistically significant. A direct comparison of the bivariate correlations between the two hierarchical regression results yielded no significant differences between the regressions, $z_r = .74$, *ns*. However, the near zero adjusted R^2, the small effect size R^2, and the near one F model statistics for the Mexican American sample suggested that a significant increase in the Mexican American sample size, and thus power, would not bring about an equivalent level of strength of association for the two groups.

DISCUSSION

This project utilized an information processing approach in conjunction with a psychophysiological measure, task-evoked pupillary dilation response, to investigate how culture may relate to cognitive ability testing. Consistent with numerous studies (reviewed in Deary & Stough, 1996; Kranzler & Jensen, 1989), a behavioral measure of visual IT, average VBM detection accuracy, was significantly related to WAIS-R (Satz-Mogel Abbreviated Procedure) (Satz & Mogel, 1962) FSIQ scores, accounting for about 17.0% of the variance after SES had been taken into account. Participants who more accurately detected VBM target stimuli had higher FSIQ scores. As predicted, a psychophysiological measure indexing processing resource allocation to the masking stimulus, average VBM mask pupillary response, significantly added to the predictions of FSIQ by uniquely accounting for 10.6% of FSIQ scores above and beyond that accounted by SES and average VBM detection accuracy. This finding is consistent with the hypothesis that individuals who wastefully allocate resources to irrelevant information, the VBM masking stimulus in this study, perform more poorly on IQ tests (e.g., Cha & Merrill, 1994; McCall, 1994; Merrill & Taube, 1996).

Consistent with the study hypotheses, detection accuracy and pupillary response measures on the VBM task were significantly related to Caucasian American students' WAIS-R scores, but did not significantly relate to Mexican American students' scores when the two groups were compatible on FSIQ. The effect size for the Mexican American sample was two fifths that of the Caucasian American sample, and the small nonsignificant effect size ($R^2 = .10$) for the Mexican American sample would require 4 times the number of Mexican American participants to be significant. This finding of differential validity meets the criteria for test bias as set forth by Cole and Moss (1989). A plausible explanation for this disassociation is that WAIS-R scores are a less valid measure of cognitive ability for Mexican American than for Caucasian American students due to some form of cultural bias on the WAIS-R that is not as much a factor on information processing and pupillary response measures. These data *do not* mean that the Mexican American sample or the Mexican American population is less intelligent than the Caucasian American sample. On the contrary, these data indicate equal cognitive ability in the Mexican American and Caucasian American samples on the theoretically more culture-fair information processing and psychophysiological measures but lower WAIS-R FSIQ test scores in Mexican Americans, possibly due to the cultural influences embedded in this test. These findings and interpretations are consistent with Fagan's (1992, 2000;

Fagan & Haiken-Vasen, 1997) theory of intelligence as processing, which may reduce cultural biases in standardized assessment. If standardized assessments such as the WAIS-R were influenced by cultural differences, then differential placement into educational and vocational programs would be expected. The problem of disproportionately high numbers of ethnic students placed in special education programs (MacMillan et al., 1993) and low numbers placed into gifted programs (Ford-Harris, 1993) might be remedied by developing assessment methods that are less influenced by culture differences. Information processing tasks and psychophysiological measures are thought to index basic cognitive processes (i.e., early processing stages of perception, discrimination, and filtering) that are minimally influenced by cultural or environmental factors.

One could consider other possible explanations for the results. A genetic hypothesis has been used to explain ethnic differences on cognitive ability testing (Jensen, 1969, 1973; Rushton, 1988). The genetic hypothesis predicts lowered performances on all cognitive assessment levels. In this study, however, there were no significant ethnic differences on the predictor variables (i.e., average VBM detection accuracy, average VBM mask pupillary response) in the samples that differed significantly on FSIQ. Differential motivation is another potential explanation for the findings in this study. However, motivational differences (Larson, Saccuzzo, & Brown, 1994) and the use of strategies (Deary & Stough, 1996) have not accounted for the IT-IQ correlation. Both groups performed equally on the information processing efficiency measures and performed at levels slightly higher on the WAIS-R than is found in the literature (Suzuki & Gutkin, 1993; Wechsler, 1981; Whitworth & Gibbons, 1986). Therefore, differential motivation is not a likely explanation for these findings. Finally, one might consider that linguistic bias could explain the findings in this study because the two groups differed on first language learned and bilingualism. However, the two groups did not differ on their performance of the Reading subtest of the WRAT3, their correlations between the WRAT3 and WAIS-R scores, or their WAIS-R VIQ scores in the FSIQ-compatible analyses. Language differences alone, therefore, cannot explain the differences found between the two groups.

The pattern of differential validity found between the two groups in this study may have clinical relevance for the assessment industry. The United States continues to increase in multicultural composition, bringing the issues regarding appropriate assessment to the forefront of professional concern. The findings in this study suggest that it may be possible in the future to use information processing and psychophysiological measures to assess cognitive ability in culturally diverse populations. There is mounting evidence that several culturally related variables are correlated with performance on currently available standardized cognitive ability tests. Accounting for SES, for example, reduces ethnic group differences (Valencia & Suzuki, 2001). Consistent with this literature, SES was associated with FSIQ in this study in the total sample. In the FSIQ-compatible samples, SES differed significantly between the two groups but had negligible impact on FSIQ for either group. The two groups also differed on several other culturally relevant variables, including first language learned, bilingualism, the number of generations the participants' families have resided in the United States, and indices of ethnic experience on the SEE. These findings are consistent with cultural cognitive ability correlates reported in the literature. For example, home environment (Valencia & Suzuki, 2001), language preference (Harris, Tulsky, & Schultheis, 2003), and length and quality of education (Ceci & Williams, 1997) have all been found to influence performance on standardized cognitive ability tests. Similarly, cross-cultural studies in neuropsychology have found associations between neuropsychological performance and bilingualism, illiteracy, level and quality of education, acculturation, and cultural variables on handedness (Ardila, 1995; Ferraro, 2002; Fletcher-Janzen, Strickland, & Reynolds, 2000). Our findings do not permit conclusions about which of these factors may have contributed to cultural bias on the WAIS-R in the Mexican American sample. Studies are needed to investigate the extent and role of these culturally related correlates with cognitive ability assessment measures.

Consistent with numerous studies spanning more than 25 years of research (reviewed in Deary & Stough, 1996; Kranzler & Jensen, 1989), a substantial correlation was found between average VBM detection accuracy and FSIQ. Most studies have used a tachistoscope to administer the task, and IT was used as the dependent variable. In contrast, a computer-administered task was used in this study and detection accuracy was the dependent variable. Despite methodological differences across studies, average VBM detection accuracy proved to be of comparable predictive value to that in previous studies.

Task-evoked pupillary response as an index of information processing efficiency proved to be a novel and powerful predictor of cognitive ability. A PCA component of pupillary response that appeared to index resource allocation to the masking stimulus (Verney et al., 2001), the average VBM mask pupillary response variable, contributed significantly to predictions in WAIS-R scores. Furthermore, the average mask pupillary response variable had the most substantial semipartial correlation in the full model regression, accounting for 11.1% of the variance in FSIQ in the presence of SES, average VBM detection accuracy, and ethnicity. Pupillary responses in a subset of these participants also significantly predicted SAT scores (Verney, Granholm, & Marshall, 2004). In a previous study (Verney et al., 2001), we found that only the pupillary responses during a late SOA condition differed significantly from the no-mask condition, suggesting that the irrelevant masking stimulus demanded additional processing resources. This earlier finding was replicated in this study. Consistent with cognitive models of selective attention (Schneider & Fisk, 1984; Shiffrin & Schneider, 1977), individuals with greater mask processing (i.e., pupillary dilation response) may have failed to inhibit the masking stimulus and thus required more resources in processing the mask before deeming it irrelevant.

A few studies have shown that wastefully allocating resources to distracting information is associated with poorer cognitive ability (Cha & Merrill, 1994; McCall, 1994; Merrill & Taube, 1996). Most of these studies were conducted on more extreme populations (e.g., individuals exhibiting mental retardation). The present study provides further support for the hypothesis that wasteful allocation of resources to irrelevant information correlates substantially with cognitive ability in a normal college student population. A few possible explanations for this finding exist. Alternative explanations may include that the mask as a separate, distinct percept in the longer SOA conditions (Michaels & Turvey, 1979, Phillips, 1974) demands more resources for individuals with lower intellectual ability (because they actively and routinely process the stimulus before determining the information to be irrelevant), that individuals with higher intellectual abilities may actively inhibit the masking stimulus before resources are allocated to processing the irrelevant information, or that individuals with higher intellectual abilities quickly automate the processing of the masking stimulus. This study was not designed to investigate these possible mechanisms (see Verney et al., 2004, for more discussion). Further research is needed to determine the role of processing irrelevant information in intelligence.

Much more research is needed on test bias and multicultural assessment (Suzuki & Valencia, 1997; Valencia et al., 2001). This study should be replicated with larger samples representing greater diversity in culture and IQ to address issues of restricted ranges in cognitive ability and ethnic experience. Other information processing tasks tapping elementary cognitive constructs may also prove useful both in predicting cognitive ability scores and honing in on the construct of intelligence. For example, working memory (e.g., Kyllonen & Christal, 1990) and attention to novel stimuli (Fagan, 1992, 2000; Fagan & Haiken-Vasen, 1997) have been found to have significant correlations with IQ tests. Similarly, other psychophysiological measurements, such as functional magnetic resonance imaging and electroencephalograms, provide valuable information regarding human information processing and the extent of cognitive activation required to perform tasks. The results found in this study suggest that combining information processing tasks with novel psychophysiological measurements may ultimately lead to more culture-fair methods of cognitive ability assessment.

REFERENCES

Ahern, S., & Beatty, J. (1979). Pupillary responses during information processing vary with scholastic aptitude test scores. *Science*, *205*, 1289-1292.

Alexander, J. R. M., & Mackenzie, B. D. (1992). Variations of the 2-line inspection time stimulus. *Personality and Individual Differences*, *13*, 1201-1211.

Ardila, A. (1995). Directions of research in cross-cultural neuropsychology. *Journal of Clinical and Experimental Neuropsychology*, *17*, 143-150.

Armour-Thomas, E. (2003). Assessment of psychometric intelligence for racial and ethnic minorities. In G. Bernal, J. E. Trimble, A. K. Burlew, & F. T. Leong (Eds.), *Handbook of racial & ethnic minority psychology* (pp. 357-374). Thousand Oaks, CA: Sage.

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, *91*, 276-292.

Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. In J. T. Cacioppo & L. G. Tassinary (Eds.), *Handbook of psychophysiology* (2nd ed., pp. 142-162). New York: Cambridge University Press.

Bosco, J. (1972). The visual information processing speed of lower- and middle-class children. *Child Development*, *43*, 1418-1422.

Breitmeyer, B. G., & Ganz, L. (1976). Implications of sustained and transient channels for theories of visual pattern masking, saccadic suppression, and information processing. *Psychological Review*, *83*, 1-36.

Brody, N. (1992). *Intelligence* (2nd ed.). New York: Harcourt, Brace, Jovanovich.

Cha, K. H., & Merrill, E. C. (1994). Facilitation and inhibition effects in visual selective attention processes of persons with and without mental retardation. *American Journal of Mental Retardation*, *98*, 594-600.

Ceci, S. J., & Williams, W. M. (1997). Schooling, intelligence, and income. *American Psychologist*, *52*, 1051-1058.

Cole, N. S., & Moss, P. A. (1989). Bias in test use. In R. L. Linn (Ed.), *Educational measurement* (3rd ed., pp. 201-219). New York: American Council on Education & Macmillan.

Davidson, J. E., & Downing, C. L. (2000). Contemporary models of intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 34-49). Cambridge, UK: Cambridge University Press.

Deary, I. J., Hunger, R., Langan, S. J., & Goodwin, G. M. (1991). Inspection time, psychometric intelligence and clinical estimates of cognitive ability in pre-senile Alzheimer's disease and Korsakoff's psychosis. *Brain*, *114*, 2543-2554.

Deary, I. J., & Stough, C. (1996). Intelligence and inspection time: Achievements, prospects, and problems. *American Psychologist*, *51*, 599-608.

Fagan, J. F. (1992). Intelligence: A theoretical viewpoint. *Current Directions in Psychological Science*, *1*, 82-86.

Fagan, J. F. (2000). A theory of intelligence as processing: Implications for society. *Psychology, Public Policy, and Law*, *6*, 168-179.

Fagan, J. F., & Haiken-Vasen, J. (1997). Selective attention to novelty as a measure of information processing across the lifespan. In J. A.

Burack & J. T. Enns (Eds.), *Attention, development, and psychopathology* (pp. 55-73). New York: Guilford.

Ferraro, F. R. (Ed.). (2002). *Minority and cross cultural aspects of neuropsychological assessment*. Lisse, Netherlands: Swets & Zeitlinger.

Fletcher-Janzen, E., Strickland, T. L., & Reynolds, C. R. (Eds.). (2000). *Handbook of cross-cultural neuropyshological assessment*. New York: Kluwer Academic/Plenum.

Ford-Harris, D. Y. (1993). The underrepresentation of minority students in gifted education: Problems and promises in recruitment and retention. *Journal of Special Education*, *32*, 4-14.

Granholm, E., Asarnow, R. F., Sarkin, A. J., & Dykes, K. L. (1996). Pupillary responses index cognitive resource limitations. *Psychophysiology*, *33*, 457-461.

Greenfield, P. M. (1997). You can't take it with you. *American Psychologist*, *52*, 1115-1124.

Haier, R. J., Siegel, B., Nuechterlein, K. H., Hazlett, M. E., Wu, T. C., Paek, J., et al. (1988). Cortical glucose metabolic rate correlates of reasoning and attention studied with positron emission tomography. *Intelligence*, *12*, 199-217.

Haier, R. J., Siegel, B., Tang, C., Abel, L., & Buchsbaum, M. S. (1992). Intelligence and changes in regional cerebral glucose metabolic rate following learning. *Intelligence*, *16*, 415-426.

Haiken-Vasen, J. (1995). *Attention to novelty and the development of implicit memory in school-aged children*. Unpublished doctoral dissertation, Case Western Reserve University, Cleveland, OH.

Harris, J. G., Tulsky, D. S., & Schultheis, M. T. (2003). Assessment of non-native English speaker: Assimilating history and research to guide clinical practice. In D. S. Tulsky & D. H. Saklofske (Eds.), *Clinical interpretation of the WAIS-III and WMS-III* (pp. 343-390). San Diego: Academic Press.

Hendrickson, A. E. (1982). The biological basis of intelligence, Part I: Theory. In H. J. Eysenck (Ed.), *A model for intelligence* (pp. 151-196). New York: Springer-Verlag.

Hendrickson, D. E. (1982). The biological basis of intelligence, Part II: Measurement. In H. J. Eysenck (Ed.), *A model for intelligence* (pp. 197-228). New York: Springer-Verlag.

Jensen, A. R. (1969). How much can we boost IQ and scholastic achievement? *Harvard Educational Review*, *39*, 1-23.

Jensen, A. R. (1973). *Educability and group differences*. London: Methuen.

Kranzler, J., & Jensen, A. R. (1989). Inspection time and intelligence: A meta-analysis. *Intelligence*, *13*, 329-347.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working memory capacity. *Intelligence*, *14*, 389-433.

Lancy, D. F. (1983). *Cross-cultural studies in cognition and mathematics*. New York: Academic Press.

Larson, G. E., Saccuzzo, D. P., & Brown, J. (1994). Motivation: Cause or confound in information processing/intelligence correlations. *Acta Psychologia*, *85*, 25-37.

Loehlin, J. C. (2000). Group differences in intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 176-193). Cambridge, UK: Cambridge University Press.

Loewenfeld, I. E. (1999). *The pupil: Anatomy, physiology, and clinical applications*. Boston: Butterworth Heinemann.

Longstreth, L. E., Walsh, D. A., Alcorn, M. D., Szeszulski, P. A., & Manis, F. R. (1986). Backward masking, IQ, SAT, and reaction time: Interrelationships and theory. *Personality and Individual Differences*, *7*, 643-651.

Lynn, R., & Holmshaw, M. (1990). Black-white differences in reaction times and intelligence. *Social Behavior and Personality*, *18*, 299-308.

MacMillan, D. L., Gresham, F. M., & Sipersein, G. N. (1993). Conceptual and psychometric concerns about the 1992 AAMR definition of mental retardation. *American Journal on Mental Retardation*, *98*, 325-335.

Malcarne, V., Chavira, D., Fernandez, S., & Liu, P-J. (2004). *The Scale of Ethnic Experience: Development and psychometric properties*. Manuscript submitted for publication.

McCall, R. (1994). What process mediates predictions of childhood IQ from infant habituation and recognition memory? Speculations on the roles of inhibition and rate of information processing. *Intelligence*, *18*, 107-125.

Merrill, E. C., & Taube, M. (1996). Negative priming and mental retardation: The process of distractor information. *American Journal of Mental Retardation*, *101*, 63-71.

Michaels, C. F., & Turvey, M. T. (1979). Central sources of visual masking: Indexing structures supporting seeing at a single, brief glance. *Psychological Research*, *41*, 1-61.

Neisser, U., Boodoo, G., Bouchard, T. J., Jr., Boykin, A. W., Brody, N., Ceci, S. J., et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, *51*, 77-101.

Nettlebeck, T., Edwards, C., & Vreugdenhil, A. (1986). Inspection time and IQ: Evidence for a mental speed-ability association. *Personality and Individual Differences*, *7*, 633-641.

Nettlebeck, T., & Rabbit, P. M. A. (1992). Aging, cognitive performance, and mental speed. *Intelligence*, *16*, 189-205.

Parks, R. W., Lowenstein, D. A., & Dondrell, K. L. (1988). Cerebral metabolic effects of a verbal fluency test. A PET scan study. *Journal of Clinical and Experimental Neuropsychology*, *10*, 565-575.

Phillips, W. A. (1974). On the distinction between sensory storage and short-term visual memory. *Perception and Psychophysics*, *16*, 283-290.

Rushton, J. P. (1988). Race differences in behavior: A review and evolutionary analysis. *Personality and Individual Differences*, *9*, 1009-1024.

Saccuzzo, D. P. (1993). Measuring individual differences in cognition in schizophrenia and other disordered states: Backward masking paradigm. In D. K. Detterman (Ed.), *Individual differences in cognition. Current topics in human intelligence* (Vol. 3, pp. 219-237). Norwood, NJ: Ablex.

Saccuzzo, D. P., Johnson, N. E., & Guertin, T. L. (1994). Information processing in gifted versus nongifted African American, Latino, Filipino, and White children: Speeded versus nonspeeded paradigms. *Intelligence*, *19*, 219-243.

Satz, P., & Mogel, S. (1962). An abbreviation of the WAIS for clinical use. *Journal of Clinical Psychology*, *18*, 77-79.

Schafer, E. W. P. (1982). Neural adaptability: A biological determinant of behavioral intelligence. *International Journal of Neuroscience*, *17*, 183-191.

Schneider, W., & Fisk, A. (1984). Automatic category search and its transfer. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *10*, 1-15.

Seagrim, G., & Lendon, R. (1980). *Furnishing the mind: A comparative study of cognitive development in Central Australian Aborigines*. Sydney, Australia: Academic Press.

Shiffrin, R. M., & Schneider, W. (1977). Controlled and automatic human information processing: II. Perceptual learning, automatic attending, and a general theory. *Psychology Review*, *84*, 127-190.

Silverstein, A. B. (1982). Validity of Satz-Mogel-Yudin-type short forms. *Journal of Consulting and Clinical Psychology*, *50*, 20-21.

Steinhauer, S. R., & Hakerem, G. (1992). The pupillary response in cognitive psychophysiology and schizophrenia. In D. Friedman & G. E. Bruder (Eds.), *Psychophysiology and experimental psychopathology: A tribute to Samuel Sutton*. New York: New York Academy of Sciences.

Suzuki, L. A., & Gutkin, T. B. (1993, August). *Racial/ethnic ability patterns on the WISC-R and theories of intelligence*. Paper presented at the 101st Annual Convention, American Psychological Association, Toronto, Canada.

Suzuki, L. A., Meller, P. J., & Ponterotto, J. G. (Eds.). (1996). *Handbook of multicultural assessment*. San Francisco: Jossey-Bass.

Suzuki, L. A., & Valencia, R. R. (1997). Race-ethnicity and measured intelligence: Educational implications. *American Psychologist*, *52*, 1103-1114.

Tabachnick, B. G., & Fidell, L. S. (1996). *Using multivariate statistics* (3rd ed.). New York: HarperCollins College Publishers.

Valencia, R. R. & Suzuki, L. A. (2001). *Intelligence testing and minority students: Foundations, performance factors, and assessment issues*. Thousand Oaks, CA: Sage.

Valencia, R. R., Suzuki, L. A., & Salinas, M. F. (2001). Test bias. In R. R. Valencia & L. A. Suzuki (Eds.), *Intelligence testing and minority students: Foundations, performance factors, and assessment issues*. Thousand Oaks, CA: Sage.

Verney, S. P., Granholm, E., & Dionisio, D. P. (2001). Pupillary response indexes cognitive processing in the visual backward masking task. *Psychophysiology*, *38*, 76-83.

Verney, S. P., Granholm, E., & Marshall, S. P. (2004). Pupillary responses on the visual backward masking task reflect general cognitive ability. *International Journal of Psychophysiology*, *52*, 23-36.

Vernon, P. A., Wickett, J. C., Bazana, P. G., & Stelmack, R. M. (2000). The neuropsychology and psychophysiology of human intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 245-264). Cambridge, UK: Cambridge University Press.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale–Revised*. New York: Psychological Corporation.

Whitworth, R. H., & Gibbons, R. T. (1986). Cross-racial comparison of the WAIS and WAIS-R. *Educational and Psychological Measurement*, *46*, 1041-1049.

Wilkinson, G. S. (1993). *WRAT3 administration manual.* Wilmington, DE: Wide Range.

Steven P. Verney, Ph.D., is an assistant professor of psychology at the University of New Mexico and a member of the NIAAA-sponsored UNM Southwest Alcohol Research Group. He completed his doctoral training at the San Diego State University (SDSU)/University of California, San Diego (UCSD) joint doctoral program in clinical psychology and his psychology postdoctoral training at the San Diego Veterans Affairs Healthcare System's (VASDHS) Mental Illness Research, Education, and Clinical Center. His research focuses on psychophysiological indices of information processing efficiency and inefficiency.

Eric Granholm, Ph.D., is professor of psychiatry at the University of California, San Diego, program director of the Schizophrenia Psychosocial Rehabilitation Program at the VA San Diego Healthcare System, and director of the UCSD/VA Psychology Internship Program. He received his Ph.D. in Clinical Psychology from the University of California, Los Angeles, in 1991. At the UCLA Neuropsychiatric Institute, he completed both a clinical psychology internship and a postdoctoral fellowship in schizophrenia. He is an active basic and clinical researcher in the areas of cognitive behavioral therapy for psychosis, cognitive neuropsychology, and psychophysiology (pupillography) in patients with schizophrenia.

Sandra P. Marshall, Ph.D., is a professor of psychology at SDSU and director of the Cognitive Ergonomics Research Facility at SDSU. For the past 15 years, she has conducted basic research sponsored by the Office of Naval Research, Air Force Office of Scientific Research, National Science Foundation, and National Institute of Education. Her projects focus on the use of eye-tracking measures to investigate aspects of cognitive processing and decision making.

Vanessa L. Malcarne, Ph.D., is currently a professor of psychology at SDSU. She is a core faculty member of the SDSU/UCSD Joint Doctoral Program in Clinical Psychology and a member of the Rebecca and John Moores UCSD Cancer Center, an NCI-designated Comprehensive Cancer Center. Dr. Malcarne's current research focuses on psychosocial interventions to improve health-related quality of life for cancer patients and survivors.

Dennis P. Saccuzzo, Ph.D., J.D., ABPP, is a professor of psychology at SDSU and an adjunct professor of law at California Western School of Law. He has received diplomate status in clinical psychology, American Board of Professional Psychology, American Board of Assessment Psychology, American Board of Forensic Examiners, and American Board of Forensic Medicine. He has served in a variety of clinical consultant activities and is the director of a practice for forensic psychology.